ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE @ DIRECT®

J. Math. Anal. Appl. 277 (2003) 303–323

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

Optimal control for semilinear evolutionary variational bilateral problem ☆

Qihong Chen

Department of Applied Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, China

Received 20 February 2002

Submitted by A. Friedman

Abstract

This paper is concerned with an optimal control problem for some semilinear evolutionary variational inequalities associated with bilateral constraints. The control domain is a general separable metric space and has no algebraic structure, in particular, it is not necessarily convex. Existence and optimality conditions of optimal pairs are established.

Keywords: Evolutionary bilateral variational inequality; Optimal control; Existence; Pontryagin's principle

1. Introduction

This paper deals with an optimal control problem in which the state y is governed by a controlled semilinear evolutionary bilateral variational inequality

$$\begin{cases} y \in W_2^{2,1}(Q) \cap L^2\big(0, T; H_0^1(\Omega)\big), \quad y|_{t=0} = y_0 \quad \text{in } \Omega, \\ \varphi \leqslant y \leqslant \psi \quad \text{in } Q, \\ \big(y_t - \Delta y - f(x, t, y, u)\big)(y - \varphi) \leqslant 0 \quad \text{in } Q, \\ \big(y_t - \Delta y - f(x, t, y, u)\big)(y - \psi) \leqslant 0 \quad \text{in } Q. \end{cases} \tag{1.1}$$

Our goal is to minimize the following cost functional:

☆ This research is supported in part by the Natural Science Foundation of China under Grant 10171059.

E-mail address: chenqih@online.sh.cn.

0022-247X/02/$ – see front matter
PII: S0022-247X(02)00569-3

$$J(y,u)=\int\limits_{Q} L\big(x,t,y(x,t),u(x,t)\big)\,dx\,dt \tag{1.2}$$

where (y,u) is a pair of state and control satisfying (1.1).

In this paper, the standard notations $Q=\Omega\times(0,T)$, $\Sigma=\partial\Omega\times(0,T)$, etc., are adopted, and

$$W_2^{2,1}(Q)=\big\{z\in L^2(Q)\mid z_t,z_x,z_{xx}\in L^2(Q)\big\},$$
$$L^2\big(0,T;H_0^1(\Omega)\big)=\left\{z:(0,T)\to H_0^1(\Omega)\;\Big|\;\int\limits_0^T\|z(\cdot,t)\|^2_{H_0^1(\Omega)}\,dt<\infty\right\}$$

are Sobolev spaces as usual.

Optimal control problems for variational inequalities have been discussed by many authors in different aspects. See [1–4,7,8,11,12,17,19,25], for examples. Some standard results for variational inequalities can be found in [10,14,20,23].

Also, the optimal control problems with state constraints have been investigated by several authors and some optimality conditions (such as Pontryagin's maximum principle, etc.) have been derived by assuming the constraint set finite codimensional (cf. [17] and references therein). Here we should note that in our problem $\varphi\leqslant y\leqslant\psi$ is a part of the state equation, not a state constraint. From this point of view, our state constraint is of whole space and therefore it is of finite codimension (actually, codimension is 0).

We point out that the above problem (i.e., state equation (1.1) with cost functional (1.2)) is different from the problem with the same cost functional (1.2) and (1.1) replaced by

$$y_t-\Delta y=f(x,t,y,u) \tag{1.3}$$

together with the state constraint

$$\varphi\leqslant y\leqslant\psi. \tag{1.4}$$

The reason is that in (1.1), on the set $\{y=\varphi\}\cup\{y=\psi\}$, y does not necessarily satisfy (1.3). However, in (1.3) and (1.4), (1.3) is required even on $\{y=\varphi\}\cup\{y=\psi\}$. Hence, the problem with (1.2)–(1.4) and that with (1.1) and (1.2) are rather different.

There are many contributions devoted to the optimal control problems for evolutionary systems. See, for examples, [2,3,16–18,24,26] and references therein, in most of which an abstract evolution equation setting was commonly used and/or the convexity of control domain was usually assumed. We note that by using the abstract framework for evolutionary systems people treat the time variable t and the spatial variable x unequally, in the sense that the variable x is "averaged" and does not appear explicitly in the whole process. And consequently, some pointwise information on the state is lost. The optimal control for semilinear parabolic equations with pointwise state constraints has been studied in [13] (and recently in [5] for boundary control problems) without using the abstract evolution equations. However, both of them have not contained the case in which the nonlinear term is multivalued. We also note that in many practical problems the control domain usually is not necessarily convex. In this paper, the framework of partial differential equation is used instead of the abstract framework, and the control domain is assumed to be a general separable metric space without algebraic structure, in particular, it is not necessarily convex.

The rest of the paper is organized as follows. In Section 2, we make some state analysis, obtain a $W_p^{2,1}$-estimate for the state, and prove the continuity of the state with respect to the control variable. Section 3 is devoted to the existence of optimal pairs. We study the approximate problems in Section 4 and derive Pontryagin's principle in Section 5.

2. State analysis

2.1. Assumptions and problem formulation

With respect to the control domain and the data involved, we make the following assumptions:

(A$_1$) $\Omega \subset \mathbb{R}^n$ is a bounded region with $C^{1,1}$ boundary $\partial\Omega$, U is a Polish space (a separable complete metric space) and

$$\mathcal{U} = \{u : Q \to U \mid u(\cdot,\cdot) \text{ is measurable}\}.$$

(A$_2$) For some $\alpha \in (0,1)$ and any $p > 1$,

$$y_0 \in C_0^\alpha(\overline{\Omega}) \cap W^{2-1/p,p}(\Omega).$$

(A$_3$) The function $f : \Omega \times [0,T] \times \mathbb{R} \times U \to \mathbb{R}$ has the following properties: $f(\cdot,\cdot,y,u)$ is measurable on $\Omega \times [0,T]$, $f(x,t,\cdot,u)$ is in $C^1(\mathbb{R})$ with $f(x,t,\cdot,\cdot)$, and $f_y(x,t,\cdot,\cdot)$ continuous on $\mathbb{R} \times U$. Moreover, there exists a constant $K > 0$, such that

$$|f_y| \leqslant K \quad \text{on } \Omega \times [0,T] \times \mathbb{R} \times U$$

and

$$\bigl|f(x,t,0,u)\bigr| \leqslant K \quad \text{on } \Omega \times [0,T] \times U.$$

(A$_4$) The function $L : \Omega \times [0,T] \times \mathbb{R} \times U \to \mathbb{R}$ satisfies the following: $L(\cdot,\cdot,y,u)$ is measurable on $\Omega \times [0,T]$, $L(x,t,\cdot,u)$ is in $C^1(\mathbb{R})$ with $L(x,t,\cdot,\cdot)$ and $L_y(x,t,\cdot,\cdot)$ continuous on $\mathbb{R} \times U$, and for any $R > 0$, there exists a constant $K_R > 0$, such that

$$|L| + |L_y| \leqslant K_R \quad \text{on } \Omega \times [0,T] \times [-R,R] \times U.$$

Let

$$W = \bigl\{y \in L^2\bigl(0,T; H_0^1(\Omega)\bigr) \bigm| y_t \in L^2\bigl(0,T; H^{-1}(\Omega)\bigr)\bigr\}.$$

By [22, Lemma 3.2, Chapter II] we know that if $y \in W$ then y is almost every where equal to a function that is continuous from $[0,T]$ into $L^2(\Omega)$. Hence, our initial condition is meaningful for any $y \in W$.

Given $\varphi, \psi \in W_p^{2,1}(Q)$ $(\forall p \geqslant 2)$ with $\varphi \leqslant \psi$ in Q, $\varphi \leqslant 0 \leqslant \psi$ on Σ and $\varphi|_{t=0} = \psi|_{t=0} = y_0$ in Ω. Set

$$\mathbf{K} = \bigl\{w \in W \mid \varphi \leqslant w \leqslant \psi \text{ a.e. in } Q \text{ and } w|_{t=0} = y_0 \text{ a.e. in } \Omega\bigr\}. \tag{2.1}$$

Clearly, $\mathbf{K}$ is a nonempty convex and closed subset of W.

If y solves (1.1), then

$$y \in \mathbf{K} \tag{2.2}$$

and, for any $w \in \mathbf{K}$, $(w-y)^+$ ($(w-y)^-$, respectively) can differ from 0 only when $y - \psi < 0$ ($y - \varphi > 0$) and therefore $y_t - \Delta y - f \geqslant 0$ ($y_t - \Delta y - f \leqslant 0$). Thus, by the divergence theorem,

$$\begin{aligned}
&\int_Q \big[y_t(w-y) + \nabla y \cdot \nabla(w-y) - f(x,t,y,u)(w-y)\big]\,dx\,dt \\
&\quad = \int_Q (y_t - \Delta y - f)(w-y)\,dx\,dt \\
&\quad = \int_Q (y_t - \Delta y - f)(w-y)^+\,dx\,dt - \int_Q (y_t - \Delta y - f)(w-y)^-\,dx\,dt \\
&\quad \geqslant 0 \quad \forall w \in \mathbf{K}.
\end{aligned} \tag{2.3}$$

On the other hand, any $y \in W_2^{2,1}(Q) \cap L^2(0,T;H_0^1(\Omega))$ satisfying (2.2) and (2.3) must be a solution of (1.1). In fact, fix any $D \subset\subset Q$ and denote by $\{\chi_n\}$ a sequence of functions from $C_c^\infty(Q)$ satisfying $0 \leqslant \chi_n \leqslant 1$, $\chi_n \to \chi_D$ (characteristic function of D) a.e. in Q, we can insert $w = y + \chi_n(\varphi - y)$ and $w = y + \chi_n(\psi - y)$ in (2.3) in turn and obtain

$$\int_Q (y_t - \Delta y - f)\chi_n(\varphi - y)\,dx\,dt \geqslant 0 \quad \text{and}$$

$$\int_Q (y_t - \Delta y - f)\chi_n(\psi - y)\,dx\,dt \geqslant 0;$$

hence also

$$\int_D (y_t - \Delta y - f)(\varphi - y)\,dx\,dt \geqslant 0 \quad \text{and} \quad \int_D (y_t - \Delta y - f)(\psi - y)\,dx\,dt \geqslant 0$$

after passing to the limit as $n \to \infty$. By the arbitrariness of D, we arrive at (1.1).

The above discussion yields a weak formulation of variational bilateral problem (1.1).

Definition 2.1. Given $u \in \mathcal{U}$. A function $y \in W$ is called a weak solution of evolutionary bilateral variational inequality (1.1), if

$$\begin{cases} y \in \mathbf{K}, \\ \int_Q \big[y_t(w-y) + \nabla y \cdot \nabla(w-y)\big]\,dx\,dt \\ \qquad \geqslant \int_Q f(x,t,y,u)(w-y)\,dx\,dt \quad \forall w \in \mathbf{K}. \end{cases} \tag{2.4}$$

Any element $u \in \mathcal{U}$ is referred to as a control. Any pair $(y,u) \in W \times \mathcal{U}$ satisfying (2.4) is called a feasible pair and the corresponding y and u will be referred to as a feasible state and a feasible control, respectively. The set of all feasible pairs is denoted by $\mathcal{F}$. It is seen that under (A_1)–(A_4) $\mathcal{U}$ coincides with the set of all feasible controls and for each $u \in \mathcal{U}$ there corresponds a unique feasible state $y \in W \cap C(\overline{Q})$ and the cost functional (1.2) is well-defined. Throughout this paper, we keep assumptions (A_1)–(A_4). Thus, we can write $J(y,u)$ as $J(u)$ without any ambiguity.

Now, we state our optimal control problem as follows:

Problem (C). Find a feasible control $\bar{u} \in \mathcal{U}$, such that

$$J(\bar{u}) = \inf_{u \in \mathcal{U}} J(u).$$

Such a $\bar{u} \in \mathcal{U}$, if it exists, is called an optimal control; the corresponding state $\bar{y}$ and the feasible pair $(\bar{y},\bar{u}) \in \mathcal{F}$ will be called an optimal state and an optimal pair, respectively.

In what follows, our main purpose is to establish the existence theorem and derive some necessary conditions for Problem (C). Our approach applies to more general cases, namely, the Laplacian in (1.1) may be replaced by a general second order elliptic operator with smooth coefficients.

2.2. $W_p^{2,1}$-estimate for state

Let us start with a $W_p^{2,1}$-estimate of state which is essential in sequel.

Proposition 2.2. *Let* (A_1)–(A_3) *hold and let* $(y,u) \in \mathcal{F}$. *Then for any* $p \geqslant 2$,

$$\|y\|_{W_p^{2,1}(Q)} \leqslant C_p, \tag{2.5}$$

where C_p is a constant independent of the control variable u.

Proof. Let us define

$$\beta(r) = \begin{cases} 0, & 0 \leqslant r < +\infty, \\ -r^2, & -1/2 \leqslant r < 0, \\ r + 1/4, & -\infty < r < -1/2, \end{cases}$$
$$\gamma(r) = \begin{cases} 0, & -\infty < r < 0, \\ r^2, & 0 \leqslant r < 1/2, \\ r - 1/4, & 1/2 \leqslant r < +\infty, \end{cases}$$

and introduce a family of approximation to the state equation (2.4):

$$\begin{cases} y_{\varepsilon t} - \Delta y_\varepsilon + \frac{1}{\varepsilon}\big[\beta(y_\varepsilon - \varphi) + \gamma(y_\varepsilon - \psi)\big] = f(x,t,y_\varepsilon,u) & \text{in } Q, \\ y_\varepsilon|_\Sigma = 0, \qquad y_\varepsilon|_{t=0} = y_0. \end{cases} \tag*{$(2.6)_\varepsilon$}$$

Obviously, for any given $u \in \mathcal{U}$ and $\varepsilon > 0$, Eq. $(2.6)_\varepsilon$ is uniquely solvable in W. The set of all pairs $(y_\varepsilon, u) \in W \times \mathcal{U}$ satisfying $(2.6)_\varepsilon$ will be denoted by $\mathcal{F}_\varepsilon$.

To obtain (2.5), it suffices to prove the following two lemmas.

Lemma 2.3. *Let* (A_1)–(A_3) *hold and let* $(y_\varepsilon, u) \in \mathcal{F}_\varepsilon$. *Then for any* $p \geqslant 2$,

$$\big\|\beta(y_\varepsilon - \varphi)\big\|_{L^p(Q)} \leqslant \varepsilon C_p, \tag{2.7}$$

$$\big\|\gamma(y_\varepsilon - \psi)\big\|_{L^p(Q)} \leqslant \varepsilon C_p, \tag{2.8}$$

and consequently,

$$\|y_\varepsilon\|_{W_p^{2,1}(Q)} \leqslant C_p, \tag{2.9}$$

where C_p *is a constant independent of* $\varepsilon > 0$ *and* $u \in \mathcal{U}$.

Proof. Define, for $r \in \mathbb{R}$, $B(r) = |\beta(r)|^{p-2}\beta(r)$ and $\Gamma(r) = |\gamma(r)|^{p-2}\gamma(r)$. Then we have

$$B(r) \leqslant 0 \quad \text{and} \quad \Gamma(r) \geqslant 0 \quad \forall r \in \mathbb{R}, \tag{2.10}$$

$$B(r) = 0 \quad \forall r \geqslant 0 \quad \text{and} \quad \Gamma(r) = 0 \quad \forall r \leqslant 0, \tag{2.11}$$

$$\begin{aligned} &B'(r) = (p-1)\big|\beta(r)\big|^{p-2}\beta'(r) \geqslant 0 \quad \text{and} \\ &\Gamma'(r) = (p-1)\big|\gamma(r)\big|^{p-2}\gamma'(r) \geqslant 0. \end{aligned} \tag{2.12}$$

Let

$$\Phi(r) = \int_0^r B(\tau)\,d\tau.$$

By (2.10)–(2.12) we further have

$$\Phi(r) \geqslant 0, \quad \Phi'(r) = \beta(r) \leqslant 0 \quad \text{in } \mathbb{R}; \qquad \Phi(r) = 0 \quad \text{in } \mathbb{R}^+.$$

Thus, we easily get

$$\begin{aligned} &\int_Q \big[(y_\varepsilon - \varphi)_t B(y_\varepsilon - \varphi) + \nabla(y_\varepsilon - \varphi)\cdot\nabla B(y_\varepsilon - \varphi)\big]\,dx\,dt \\ &\quad = \int_\Omega \Phi(y_\varepsilon - \varphi)|_{t=T}\,dx + \int_Q B'(y_\varepsilon - \varphi)\big|\nabla(y_\varepsilon - \varphi)\big|^2\,dx\,dt \geqslant 0. \end{aligned} \tag{2.13}$$

By (A_3), using a change of variable $Y = e^{-Kt}y$ if necessary, we may assume without loss of generality that

$$f_y \leqslant 0. \tag{2.14}$$

This, together with (2.10) and (2.11), yields

$$\int_Q f(x,t,y_\varepsilon,u)B(y_\varepsilon - \varphi)\,dx\,dt \leqslant \int_Q f(x,t,\varphi,u)B(y_\varepsilon - \varphi)\,dx\,dt. \tag{2.15}$$

Multiplying $(2.6)_\varepsilon$ by $\varepsilon B(y_\varepsilon - \varphi)$ and integrating over Q, noting also that (2.11) implies

$$\beta(y_\varepsilon - \varphi)\gamma(y_\varepsilon - \psi) = 0 \quad \text{a.e. in } Q,$$

we obtain

$$\varepsilon\int_Q \bigl[y_{\varepsilon t}B(y_\varepsilon-\varphi)+\nabla y_\varepsilon\cdot\nabla B(y_\varepsilon-\varphi)\bigr]\,dx\,dt+\int_Q \bigl|\beta(y_\varepsilon-\varphi)\bigr|^p\,dx\,dt$$
$$=\varepsilon\int_Q f(x,t,y_\varepsilon,u)B(y_\varepsilon-\varphi)\,dx\,dt. \tag{2.16}$$

Then, using (2.13), (2.15) and Hölder's inequality, we can deduce from (2.16) that

$$\begin{aligned}&\bigl\|\beta(y_\varepsilon-\varphi)\bigr\|^p_{L^p(Q)}\\ &\quad\leqslant\varepsilon\int_Q\bigl\{f(x,t,\varphi,u)B(y_\varepsilon-\varphi)-\bigl[\varphi_t B(y_\varepsilon-\varphi)+\nabla\varphi\cdot\nabla B(y_\varepsilon-\varphi)\bigr]\bigr\}\,dx\,dt\\ &\quad=\varepsilon\int_Q\bigl[f(x,t,\varphi,u)-\varphi_t+\Delta\varphi\bigr]B(y_\varepsilon-\varphi)\,dx\,dt\\ &\quad\leqslant\varepsilon\bigl[\bigl\|f\bigl(\cdot,\cdot,\varphi(\cdot,\cdot),u(\cdot,\cdot)\bigr)\bigr\|_{L^p(Q)}+\|\varphi\|_{W_p^{2,1}(Q)}\bigr]\bigl\|\beta(y_\varepsilon-\varphi)\bigr\|^{p-1}_{L^p(Q)}.\end{aligned}$$

By (A_3) we know that the function $f(\cdot,\cdot,\varphi(\cdot,\cdot),u(\cdot,\cdot))$ has a uniform bound independent of u. This means

$$\bigl\|f\bigl(\cdot,\cdot,\varphi(\cdot,\cdot),u(\cdot,\cdot)\bigr)\bigr\|_{L^p(Q)}\leqslant C_p$$

with C_p being independent of $\varepsilon>0$ and $u\in\mathcal{U}$. Thus, (2.7) follows.

The estimate (2.8) can be obtained similarly, and (2.9) follows immediately from (2.7), (2.8), and the standard parabolic L^P-estimate (cf. [15]). □

Lemma 2.4. *Let* (A_1)–(A_3) *hold and let* $(y_\varepsilon,u)\in\mathcal{F}_\varepsilon$, $(y,u)\in\mathcal{F}$. *Then, for any* $p>(n+2)/2$, *as* $\varepsilon\to 0$,

$$y_\varepsilon\to y\quad\textit{weakly in } W_p^{2,1}(Q)\textit{ and strongly in } C^{\theta,\theta/2}(\overline{Q})\cap L^2\bigl(0,T;H_0^1(\Omega)\bigr)$$

for some $\theta\in(0,1)$.

Proof. First, by Simon's compactness lemma (cf. [21]), we know that if $p>(n+2)/2$ then any bounded subset of $W_p^{2,1}(Q)$ is compact in $C^{\theta,\theta/2}(\overline{Q})\cap L^2(0,T;H_0^1(\Omega))$ for some $\theta\in(0,1)$. Thus, by virtue of estimate (2.9), we can extract a subsequence (still denoted by itself), such that

$$y_\varepsilon\to y^*\quad\text{weakly in } W_p^{2,1}(Q)\text{ and strongly in } C^{\theta,\theta/2}(\overline{Q})\cap L^2\bigl(0,T;H_0^1(\Omega)\bigr)$$

for some $\theta\in(0,1)$.

For any $\eta\in L^2(0,T;H_0^1(\Omega))$ with $\eta\geqslant 0$ a.e. in Q, it follows from $(2.6)_\varepsilon$ that

$$\int_Q\bigl[\beta(y_\varepsilon-\varphi)+\gamma(y_\varepsilon-\psi)\bigr]\eta\,dx\,dt$$
$$=\varepsilon\int_Q\bigl[f(x,t,y_\varepsilon,u)\eta-y_{\varepsilon t}\eta-\nabla y_\varepsilon\cdot\nabla\eta\bigr]\,dx\,dt\to 0.$$

Then, by Lebesgue's dominated convergence theorem, we have

$$\int_Q \big[\beta(y^* - \varphi) + \gamma(y^* - \psi)\big]\eta\, dx\, dt = 0$$
$$\forall \eta \in L^2\big(0, T; H_0^1(\Omega)\big) \text{ with } \eta \geqslant 0 \text{ a.e. in } Q.$$

This implies that

$$\beta(y^* - \varphi) + \gamma(y^* - \psi) = 0 \quad \text{a.e. in } Q$$

and, by the definition of $\beta(\cdot)$ and $\gamma(\cdot)$,

$$\varphi(x, t) \leqslant y^*(x, t) \leqslant \psi(x, t) \quad \text{a.e. in } Q.$$

Clearly, $y^*|_{t=0} = y_0$. Hence, $y^* \in \mathbf{K}$.

For any $w \in \mathbf{K}$, since $\beta(y_\varepsilon - \varphi)$ can differ from 0 only when $y_\varepsilon < \varphi \leqslant w$ and $\gamma(y_\varepsilon - \psi)$ can differ from 0 only when $y_\varepsilon > \psi \geqslant w$, we deduce from $(2.6)_\varepsilon$ that

$$\begin{aligned}
&\int_Q \big[y_{\varepsilon t}(w - y_\varepsilon) + \nabla y_\varepsilon \cdot \nabla(w - y_\varepsilon)\big]\, dx\, dt \\
&\quad = -\frac{1}{\varepsilon}\int_Q \big[\beta(y_\varepsilon - \varphi) + \gamma(y_\varepsilon - \psi)\big](w - y_\varepsilon)\, dx\, dt \\
&\qquad + \int_Q f(x, t, y_\varepsilon, u)(w - y_\varepsilon)\, dx\, dt \\
&\quad \geqslant \int_Q f(x, t, y_\varepsilon, u)(w - y_\varepsilon)\, dx\, dt.
\end{aligned} \tag{2.17}$$

Taking the limit in (2.17) we see that y^* is a weak solution of (2.4).

By uniqueness we must have that $y^* = y$ and the whole sequence $\{y_\varepsilon\}$ converges to y. □

2.3. Continuous dependence of state on control

Recall that, in the control set $\mathcal{U}$, Ekeland's distance is defined as

$$d(u, v) = m\big(\{(x, t) \in Q \mid u(x, t) \neq v(x, t)\}\big) \quad \forall u, v \in \mathcal{U}$$

where m denotes the Lebesgue measure on $\mathbb{R}^{n+1}$. It is easy to check that $(\mathcal{U}, d)$ is a complete metric space.

The following result gives the continuity of the state y with respect to the control u under the above metric.

Proposition 2.5. *Let* (A_1)–(A_3) *hold and let* $(y, u), (y_k, u_k) \in \mathcal{F}$ $(k = 1, 2, \ldots)$. *If* $d(u_k, u) \to 0$, *then*

$$\lim_{k\to\infty} \|y_k - y\|_{C^{\theta,\theta/2}(\overline{Q}) \cap L^2(0,T;H_0^1(\Omega))} = 0 \quad \big(\textit{for some } \theta \in (0, 1)\big).$$

Proof. From Proposition 2.2 and Simon's compactness lemma, we conclude that, for some subsequence,

$$y_k \to y^* \quad \text{weakly in } W_p^{2,1}(Q) \text{ and strongly in } C^{\theta,\theta/2}(\overline{Q}) \cap L^2\big(0,T;H_0^1(\Omega)\big)$$

for some $\theta \in (0,1)$. Clearly,

$$y^*|_{t=0} = y_0 \tag{2.18}$$

and

$$\varphi(x,t) \leqslant y^*(x,t) \leqslant \psi(x,t) \quad \text{a.e. in } Q. \tag{2.19}$$

Note that

$$\begin{aligned}
&\big\| f\big(\cdot,\cdot,y_k(\cdot,\cdot),u_k(\cdot,\cdot)\big) - f\big(\cdot,\cdot,y^*(\cdot,\cdot),u(\cdot,\cdot)\big)\big\|_{L^2(Q)} \\
&\quad \leqslant \big\| f\big(\cdot,\cdot,y_k(\cdot,\cdot),u_k(\cdot,\cdot)\big) - f\big(\cdot,\cdot,y^*(\cdot,\cdot),u_k(\cdot,\cdot)\big)\big\|_{L^2(Q)} \\
&\qquad + \big\| f\big(\cdot,\cdot,y^*(\cdot,\cdot),u_k(\cdot,\cdot)\big) - f\big(\cdot,\cdot,y^*(\cdot,\cdot),u(\cdot,\cdot)\big)\big\|_{L^2(Q)} \\
&\quad \leqslant C\big\{\|y_k - y^*\|_{L^2(Q)} + d(u_k,u)^{1/2}\big\} \to 0.
\end{aligned}$$

Passing to the limit in (2.4), in which u and y are replaced by u_k and y_k, respectively, we obtain

$$\int_Q \big[y_t^*(w-y^*) + \nabla y^* \cdot \nabla(w-y^*)\big]\,dx\,dt \geqslant \int_Q f(x,t,y^*,u)(w-y^*)\,dx\,dt \qquad \forall w \in \mathbf{K}.$$

This, combined with (2.18) and (2.19), means that y^* is a solution of (2.4).

Finally, the uniqueness ensures that $y^* = y$ and the convergence of the whole state sequence $\{y_k\}$. □

The following is a direct consequence of the above result:

Corollary 2.6. *Let* (A_1)–(A_4) *hold. Then* $J(u)$ *is continuous on* $(\mathcal{U}, d)$.

3. Existence

3.1. Cesari property and measurable selection theorem

Let us first recall the following

Definition 3.1 (cf. [6,17]). Let Y be a Banach space and Z be a metric space. We say a multifunction $\Lambda : Z \to 2^Y$ has the Cesari property at $z \in Z$ if

$$\bigcap_{\delta>0} \overline{\mathrm{co}}\,\Lambda\big(O_\delta(z)\big) = \Lambda(z),$$

where $\overline{\text{co}}\, E$ stands for the closed convex hull of the set E and $O_\delta(z)$ is the δ-neighborhood of the point z. If Λ has the Cesari property at every point $z \in Z$, we simply say that Λ has the Cesari property on Z.

Definition 3.2. Let $D \subset \mathbb{R}^d$ be some Lebesgue measurable set and U be a Polish space. Let $M : D \to 2^U$ be a multifunction. Function $u : D \to U$ is called a selection of $M(\cdot)$ if

$$u(s) \in M(s) \quad \text{a.e. } s \in D.$$

If such a u is measurable, then u is called a measurable selection of $M(\cdot)$.

The following gives the existence of measurable selections.

Lemma 3.3. *Let $M : D \to 2^U$ be measurable taking closed set values. Then $M(\cdot)$ admits a measurable selection.*

We refer the readers to [17, pp. 100, 101] for the proof of Lemma 3.3.

3.2. Existence of optimal controls

To establish the existence for Problem (C), we first introduce the following set:

$$\Lambda(x,t,y) = \bigl\{(\xi,\eta) \in \mathbb{R}^2 \mid \xi \geqslant L(x,t,y,u),\ \eta = f(x,t,y,u), u \in U\bigr\}$$

and make the following assumption:

(A_5) For almost all $(x,t) \in Q$, the mapping $y \mapsto \Lambda(x,t,y)$ has the Cesari property on $\mathbb{R}$.

Then we have

Theorem 3.4 (Existence theorem). *Let* (A_1)–(A_4) *and* (A_5) *hold. Then Problem* (C) *admits at least one optimal control $\bar{u} \in \mathcal{U}$.*

Proof. Let $\{u_k\} \subset \mathcal{U}$ be a minimizing sequence satisfying

$$J(u_k) \leqslant \inf_{u\in\mathcal{U}} J(u) + \frac{1}{k}. \tag{3.1}$$

Take $p > \max\{(n+2)/2, 2\}$. By Proposition 2.2 we know that the corresponding state y_k satisfies

$$\|y_k\|_{W_p^{2,1}(Q)} \leqslant C_p \tag{3.2}$$

with C_p independent of k. Thus, we may let, extracting some subsequence if necessary,

$$y_k \to \bar{y} \quad \text{weakly in } W_p^{2,1}(Q) \text{ and strongly in } C^{\theta,\theta/2}(\overline{Q}) \cap L^2\bigl(0,T;H_0^1(\Omega)\bigr) \tag{3.3}$$

for some $\theta \in (0,1)$ and some $\bar{y} \in W_p^{2,1}(Q) \cap L^2(0,T;H_0^1(\Omega))$.

By (3.2) and (A_3), the function $f(\cdot,\cdot,y_k(\cdot,\cdot),u_k(\cdot,\cdot))$ is uniformly bounded. Hence we may further assume

$$f\big(\cdot,\cdot,y_k(\cdot,\cdot),u_k(\cdot,\cdot)\big)\to\overline{f}(\cdot,\cdot)\quad\text{weakly in } L^p(Q)\tag{3.4}$$

for some $\overline{f}\in L^\infty(Q)$. Then, by Mazur theorem, we can find $\alpha_{ij}\geqslant 0$, $\sum_{i\geqslant 1}\alpha_{ij}=1$ $\forall j$, such that

$$\eta_j(\cdot,\cdot)=\sum_{i\geqslant 1}\alpha_{ij}f\big(\cdot,\cdot,y_{i+j}(\cdot,\cdot),u_{i+j}(\cdot,\cdot)\big)\to\overline{f}(\cdot,\cdot)\quad\text{strongly in } L^p(Q).$$

Set

$$\xi_j(\cdot,\cdot)=\sum_{i\geqslant 1}\alpha_{ij}L\big(\cdot,\cdot,y_{i+j}(\cdot,\cdot),u_{i+j}(\cdot,\cdot)\big)\tag{3.5}$$

and

$$\overline{L}(x,t)=\varliminf_{j\to\infty}\xi_j(x,t)\quad\text{a.e. } (x,t)\in Q.\tag{3.6}$$

The convergence (3.3) implies that, for any $\delta>0$, there exists a j_0, such that for $j\geqslant j_0$,

$$\big(\xi_j(x,t),\eta_j(x,t)\big)\in\operatorname{co}\Lambda\big(x,t,O_\delta\big(\overline{y}(x,t)\big)\big)\quad\text{a.e. } (x,t)\in Q.$$

Thus, for any $\delta>0$, we have

$$\big(\overline{L}(x,t),\overline{f}(x,t)\big)\in\overline{\operatorname{co}}\Lambda\big(x,t,O_\delta\big(\overline{y}(x,t)\big)\big)\quad\text{a.e. } (x,t)\in Q$$

and then, by (A_5),

$$\big(\overline{L}(x,t),\overline{f}(x,t)\big)\in\Lambda\big(x,t,\overline{y}(x,t)\big)\quad\text{a.e. } (x,t)\in Q.$$

Now consider a multifunction $M:Q\to 2^U$ defined as follows:

$$M(x,t)=\big\{u\in U\ \big|\ \overline{L}(x,t)\geqslant L\big(x,t,\overline{y}(x,t),u\big),\ \overline{f}(x,t)=f\big(x,t,\overline{y}(x,t),u\big)\big\}$$
$$(x,t)\in Q.$$

By (A_3) and (A_4) we see that M is closed set valued. Then, making use of Lemma 3.3, we can find a $\overline{u}\in\mathcal{U}$ such that

$$\begin{cases}\overline{L}(x,t)\geqslant L\big(x,t,\overline{y}(x,t),\overline{u}(x,t)\big),\\ \overline{f}(x,t)=f\big(x,t,\overline{y}(x,t),\overline{u}(x,t)\big)\quad\text{a.e. } (x,t)\in Q.\end{cases}\tag{3.7}$$

We claim that $\overline{y}$ is the state corresponding to $\overline{u}$, i.e.,

$$(\overline{y},\overline{u})\in\mathcal{F}.\tag{3.8}$$

In fact, since $(y_k,u_k)\in\mathcal{F}$, from the convergence (3.3) we have

$$\overline{y}|_{t=0}=y_0\tag{3.9}$$

and

$$\varphi(x,t)\leqslant\overline{y}(x,t)\leqslant\psi(x,t)\quad\text{a.e. } (x,t)\in Q.\tag{3.10}$$

Moreover, for any k, the feasibility of (y_k,u_k) gives

$$\int_Q \big[y_{kt}(w-y_k)+\nabla y_k\cdot\nabla(w-y_k)\big]\,dx\,dt$$
$$\geqslant \int_Q f(x,t,y_k,u_k)(w-y_k)\,dx\,dt\quad \forall w\in\mathbf{K},$$

and the convergence (3.3), (3.4), combined with (3.7) yields

$$\int_Q \big[\bar y_t(w-\bar y)+\nabla\bar y\cdot\nabla(w-\bar y)\big]\,dx\,dt$$
$$\geqslant \int_Q f(x,t,\bar y,\bar u)(w-\bar y)\,dx\,dt\quad \forall w\in\mathbf{K}. \tag{3.11}$$

Thus, by (3.9)–(3.11), the feasibility (3.8) is verified.

Finally, we can deduce from (3.7), (3.6), (3.5), (3.1) and Fatou's lemma, that

$$\begin{aligned}
J(\bar u) &= \int_Q L\big(x,t,\bar y(x,t),\bar u(x,t)\big)\,dx\,dt\\
&\leqslant \int_Q \overline{L}(x,t)\,dx\,dt=\int_Q \varliminf_{j\to\infty}\xi_j(x,t)\,dx\,dt\\
&\leqslant \varliminf_{j\to\infty}\int_Q \xi_j(x,t)\,dx\,dt=\varliminf_{j\to\infty}\sum_{i\geqslant 1}\alpha_{ij}J(u_{i+j})\\
&\leqslant \varliminf_{j\to\infty}\sum_{i\geqslant 1}\alpha_{ij}\left(\inf_{u\in\mathcal{U}}J(u)+\frac{1}{j}\right)=\varliminf_{j\to\infty}\left(\inf_{u\in\mathcal{U}}J(u)+\frac{1}{j}\right)\\
&=\inf_{u\in\mathcal{U}}J(u).
\end{aligned}$$

Hence, $\bar u$ is an optimal control of Problem (C). □

4. Analysis of approximation

4.1. Continuity of approximate functional

Let us consider the following approximate functional:

$$J_\varepsilon(u)=\int_Q L\big(x,t,y_\varepsilon(x,t),u(x,t)\big)\,dx\,dt \tag{4.1}$$

where $(y_\varepsilon,u)\in\mathcal{F}_\varepsilon$.

The following result gives the continuity of $J_\varepsilon(\cdot)$ on $(\mathcal{U},d)$ and the convergence of J_ε as $\varepsilon\to 0$.

Proposition 4.1. *Let* (A_1)–(A_4) *hold. Then we have the following*:

(i) *For any fixed* $\varepsilon > 0$, $J_\varepsilon(u)$ *is continuous on* $(\mathcal{U}, d)$;
(ii) *For any given* $u \in \mathcal{U}$, $\lim_{\varepsilon\to 0} J_\varepsilon(u) = J(u)$.

Proof. (i) It suffices to prove the continuous dependence of y_ε on u. Let (y_ε, u), $(y_{\varepsilon,k}, u_k) \in \mathcal{F}_\varepsilon$ $(k = 1, 2, \ldots)$ with $d(u_k, u) \to 0$. Applying the parabolic L^p-estimate to the equation satisfied by the difference $y_{\varepsilon,k} - y_\varepsilon$, we can obtain

$$\begin{aligned}\|y_{\varepsilon,k} - y_\varepsilon\|_{W_p^{2,1}(Q)} &\leqslant C\big\| f\big(\cdot,\cdot, y_\varepsilon(\cdot,\cdot), u_k(\cdot,\cdot)\big) - f\big(\cdot,\cdot, y_\varepsilon(\cdot,\cdot), u(\cdot,\cdot)\big)\big\|_{L^p(Q)} \\ &\leqslant C d(u_k, u)^{1/p}.\end{aligned}$$

Thus, we have

$$\lim_{k\to\infty} \|y_{\varepsilon,k} - y_\varepsilon\|_{W_p^{2,1}(Q)} = 0.$$

(ii) is immediately obtained from Lemma 2.4 and (A_4). □

4.2. *Variation of approximate state and functional*

Since the control domain U is merely a metric space and there is no convexity in general, only the spike perturbation of the control is allowed when we derive the necessary conditions. Note that for each $(y_\varepsilon, u) \in \mathcal{F}_\varepsilon$, y_ε solves semilinear parabolic equation $(2.6)_\varepsilon$. Thus, using an argument similar to that in [13], we can present the following "Taylor expansion" formula for the approximate state y_ε and functional $J_\varepsilon(u)$.

Proposition 4.2. *Let* (A_1)–(A_4) *hold and let* $(y_\varepsilon, u) \in \mathcal{F}_\varepsilon$ *and* $v \in \mathcal{U}$ *be fixed. Then, for any* $\rho \in (0, 1)$, *there exists a measurable set* $E^\rho \subset Q$ *with* $m(E^\rho) = \rho m(Q)$ *such that if we define* $u^\rho \in \mathcal{U}$ *by*

$$u^\rho(x,t) = \begin{cases} u(x,t) & \text{if } (x,t) \in Q \setminus E^\rho, \\ v(x,t) & \text{if } (x,t) \in E^\rho, \end{cases}$$

and let $(y_\varepsilon^\rho, u^\rho) \in \mathcal{F}_\varepsilon$, *then there holds*

$$\begin{cases} y_\varepsilon^\rho = y_\varepsilon + \rho z + r^\rho, \\ \lim_{\rho\to 0} \frac{1}{\rho}\|r^\rho\|_{C^{\theta,\theta/2}(\overline{Q})} = 0 \end{cases}$$

for some $\theta \in (0, 1)$, *and*

$$\begin{cases} J_\varepsilon(u^\rho) = J_\varepsilon(u) + \rho j + e^\rho, \\ \lim_{\rho\to 0} \frac{1}{\rho}|e^\rho| = 0 \end{cases}$$

where z *and* j *satisfy the following*:

$$\begin{cases} z_t - \Delta z + \big\{\frac{1}{\varepsilon}\big[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\big] - f_y(x,t,y_\varepsilon,u)\big\} z \\ \quad = f(x,t,y_\varepsilon,v) - f(x,t,y_\varepsilon,u) \quad \text{in } Q, \\ z|_{\partial_p Q} = 0, \end{cases}$$

and

$$j = \int_Q \big[L_y(x,t,y_\varepsilon,u)z + L(x,t,y_\varepsilon,v) - L(x,t,y_\varepsilon,u)\big]\,dx\,dt.$$

In the above, $\partial_p Q = \Sigma \cup (\Omega \times \{0\})$ is the so called parabolic boundary of Q.

4.3. Convergence of approximate state

This subsection is devoted to some convergence results for approximate state which can be regarded as some improvements of Lemma 2.4.

Proposition 4.3. *Let* (A_1)–(A_3) *hold and let* $\{u_\varepsilon\} \subset \mathcal{U}$ *be any sequence,* $(y_\varepsilon, u_\varepsilon) \in \mathcal{F}_\varepsilon$ *and* $(y^\varepsilon, u_\varepsilon) \in \mathcal{F}$. *Then*

$$\lim_{\varepsilon\to 0} \|y_\varepsilon - y^\varepsilon\|_{C^{\theta,\theta/2}(\overline{Q})\cap L^2(0,T;H_0^1(\Omega))} = 0 \tag{4.2}$$

for some $\theta \in (0,1)$.

Proof. By Proposition 2.2 and Lemma 2.3, we have that, for any $p \geqslant 2$,

$$\|y_\varepsilon\|_{W_p^{2,1}(Q)} + \|y^\varepsilon\|_{W_p^{2,1}(Q)} \leqslant C_p$$

with C_p independent of $\varepsilon > 0$. Thus, recalling Simon's compactness lemma, we may assume that, for some subsequence and some $\gamma \in (0,1)$,

$$y_\varepsilon \to y \quad \text{strongly in } C^{\gamma,\gamma/2}(\overline{Q}) \cap L^2\big(0,T;H_0^1(\Omega)\big)$$

and

$$\|y^\varepsilon\|_{W_2^{2,1}(Q)\cap C^{\gamma,\gamma/2}(\overline{Q})} \leqslant C \quad \text{with } C \text{ independent of } \varepsilon > 0. \tag{4.3}$$

An argument similar to that in the proof of Lemma 2.4 yields

$$\varphi(x,t) \leqslant y(x,t) \leqslant \psi(x,t) \quad \text{a.e. in } Q.$$

Then, letting $z_\varepsilon = y_\varepsilon \vee \varphi \wedge \psi$, we have

$$z_\varepsilon \to y \quad \text{strongly in } L^2\big(0,T;H_0^1(\Omega)\big)$$

and, consequently,

$$\|z_\varepsilon - y_\varepsilon\|_{L^2(0,T;H_0^1(\Omega))} \to 0. \tag{4.4}$$

Recalling that y_ε and y^ε solve $(2.6)_\varepsilon$ and (2.4), respectively, we have

$$\begin{aligned}
&\int_Q \bigg\{ y_{\varepsilon t}(y_\varepsilon - y^\varepsilon) + \nabla y_\varepsilon \cdot \nabla (y_\varepsilon - y^\varepsilon) \\
&\qquad + \frac{1}{\varepsilon}\big[\beta(y_\varepsilon - \varphi) + \gamma(y_\varepsilon - \psi)\big](y_\varepsilon - y^\varepsilon) \bigg\}\,dx\,dt \\
&\quad = \int_Q f(x,t,y_\varepsilon,u_\varepsilon)(y_\varepsilon - y^\varepsilon)\,dx\,dt
\end{aligned} \tag{4.5}$$

and

$$\int_Q \big[y_t^\varepsilon(z_\varepsilon - y^\varepsilon) + \nabla y^\varepsilon \cdot \nabla(z_\varepsilon - y^\varepsilon)\big]\,dx\,dt$$
$$\geqslant \int_Q f(x,t,y^\varepsilon,u_\varepsilon)(z_\varepsilon - y^\varepsilon)\,dx\,dt \tag{4.6}$$

(noting that $z_\varepsilon \in \mathbf{K}$).

By the monotonicity of $\beta(\cdot)$, $\gamma(\cdot)$ and $f(x,t,\cdot,u)$ (cf. (2.14)), we see that

$$\int_Q \big[f(x,t,y_\varepsilon,u_\varepsilon) - f(x,t,y^\varepsilon,u_\varepsilon)\big](y_\varepsilon - y^\varepsilon)\,dx\,dt \leqslant 0, \tag{4.7}$$

$$\int_Q \beta(y_\varepsilon - \varphi)(y_\varepsilon - y^\varepsilon)\,dx\,dt \geqslant 0, \tag{4.8}$$

and

$$\int_Q \gamma(y_\varepsilon - \psi)(y_\varepsilon - y^\varepsilon)\,dx\,dt \geqslant 0. \tag{4.9}$$

Here we have used the fact that $y^\varepsilon \geqslant \varphi > y_\varepsilon$ when $y_\varepsilon < \varphi$ and $y^\varepsilon \leqslant \psi < y_\varepsilon$ when $y_\varepsilon > \psi$.

From (4.3)–(4.9) we may deduce that

$$\int_Q \big|\nabla(y_\varepsilon - y^\varepsilon)\big|^2\,dx\,dt$$
$$\leqslant \int_Q \big[y_t^\varepsilon(z_\varepsilon - y_\varepsilon) + \nabla y^\varepsilon \cdot \nabla(z_\varepsilon - y_\varepsilon) - f(x,t,y^\varepsilon,u_\varepsilon)(z_\varepsilon - y_\varepsilon)\big]\,dx\,dt$$
$$\leqslant C\|z_\varepsilon - y_\varepsilon\|_{L^2(0,T;H_0^1(\Omega))} \to 0$$

and hence, by an interpolation result (see Lemma 4.4 below), we have

$$\|y_\varepsilon - y^\varepsilon\|_{C^{\theta,\theta/2}(\overline{Q})} \to 0$$

for some $0 < \theta < \gamma < 1$. □

Lemma 4.4. *Suppose that $\partial\Omega$ is Lipschitz continuous. Then there exists a constant C, depending only on Ω and T such that, for any $0 \leqslant \theta < \gamma < 1$ and $0 < p \leqslant \infty$,*

$$\|\eta\|_{C^{\theta,\theta/2}(\overline{Q})} \leqslant 4\delta\|\eta\|_{C^{\gamma,\gamma/2}(\overline{Q})} + \frac{3C^{1/p}}{\delta^\kappa}\|\eta\|_{L^p(\overline{Q})} \quad \forall 0 < \delta \leqslant 1,\ \eta \in C^{\gamma,\gamma/2}(\overline{Q})$$

where

$$\kappa = \frac{\theta}{\gamma - \theta} + \frac{n+2}{(\gamma - \theta)p}.$$

We refer the readers to [13] for its proof.

Proposition 4.3, combined with Proposition 2.5, yields

Proposition 4.5. *Let* (A_1)–(A_3) *hold and let* $\{u_\varepsilon\} \subset \mathcal{U}$ *be any sequence,* $u \in \mathcal{U}$*,* $(y_\varepsilon, u_\varepsilon) \in \mathcal{F}_\varepsilon$*, and* $(y, u) \in \mathcal{F}$*. If* $d(u_\varepsilon, u) \to 0$*, then*

$$\lim_{\varepsilon \to 0} \|y_\varepsilon - y\|_{C^{\theta,\theta/2}(\overline{Q}) \cap L^2(0,T;H_0^1(\Omega))} = 0$$

for some $\theta \in (0, 1)$.

4.4. Convergence theorem

Now we can establish the following convergence theorem which is crucial in applying Ekeland's variational principle later.

Theorem 4.6. *Let* (A_1)–(A_4) *hold and denote*

$$\overline{J}_\varepsilon = \inf_{u \in \mathcal{U}} J_\varepsilon(u), \qquad \overline{J} = \inf_{u \in \mathcal{U}} J(u). \tag{4.10}$$

Then

$$\lim_{\varepsilon \to 0} \overline{J}_\varepsilon = \overline{J}. \tag{4.11}$$

To prove Theorem 4.6 we need the following lemma which can be easily obtained from Proposition 4.3.

Lemma 4.7. *Let* (A_1)–(A_4) *hold. Then, for any sequence* $\{u_\varepsilon\} \subset \mathcal{U}$*,*

$$\lim_{\varepsilon \to 0} \big[J_\varepsilon(u_\varepsilon) - J(u_\varepsilon)\big] = 0.$$

Proof of Theorem 4.6. For any $\varepsilon > 0$ one can find $u_\varepsilon \in \mathcal{U}$, such that

$$J_\varepsilon(u_\varepsilon) < \overline{J}_\varepsilon + \varepsilon.$$

Then, by Lemma 4.7, we have

$$\varliminf_{\varepsilon \to 0} \overline{J}_\varepsilon \geqslant \varliminf_{\varepsilon \to 0} J_\varepsilon(u_\varepsilon) = \varliminf_{\varepsilon \to 0} \big[J(u_\varepsilon) + J_\varepsilon(u_\varepsilon) - J(u_\varepsilon)\big] = \varliminf_{\varepsilon \to 0} J(u_\varepsilon) \geqslant \overline{J}. \tag{4.12}$$

On the other hand, let $u_\delta \in \mathcal{U}$ be such that

$$J(u_\delta) < \overline{J} + \delta.$$

Then, using Lemma 4.7 again, we have

$$\varlimsup_{\delta \to 0} \overline{J}_\delta \leqslant \varlimsup_{\delta \to 0} J_\delta(u_\delta) = \varlimsup_{\delta \to 0} \big[J(u_\delta) + J_\delta(u_\delta) - J(u_\delta)\big] = \varlimsup_{\delta \to 0} J(u_\delta) \leqslant \overline{J}. \tag{4.13}$$

Hence, (4.11) follows from (4.12) and (4.13). □

5. Pontryagin's principle

Now we are in a position to derive Pontryagin's principle for Problem (C) by applying Ekeland's variational principle which is reviewed below for convenience' sake.

Theorem 5.1 (Ekeland's variational principle, cf. [9]). *Let (V, d) be a complete metric space and $J : V \to (-\infty, +\infty]$ be a proper lower semicontinuous function bounded from below. Let $\alpha > 0$, $\bar{v} \in V$ such that*

$$J(\bar{v}) \leqslant \inf_{v \in V} J(v) + \alpha^2.$$

Then there exists a $v_\alpha \in V$ satisfying

$$J(v_\alpha) \leqslant J(\bar{v}), \qquad d(v_\alpha, \bar{v}) \leqslant \alpha,$$
$$-\alpha d(v, v_\alpha) \leqslant J(v) - J(v_\alpha) \quad \forall v \in V.$$

Theorem 5.2 (Pontryagin's principle). *Let* (A_1)–(A_4) *hold and $(\bar{y}, \bar{u}) \in \mathcal{F}$ be an optimal pair for Problem* (C). *Then there exist $\bar{p} \in L^2(0, T; H_0^1(\Omega))$ and $\bar{\mu} \in \mathcal{M}_0(\overline{Q})$ satisfying* (*in some weak sense*)

$$\begin{cases} -\bar{p}_t - \Delta\bar{p} - f_y(x,t,\bar{y},\bar{u})\bar{p} = L_y(x,t,\bar{y},\bar{u}) - \bar{\mu} & \text{in } Q, \\ \bar{p}|_\Sigma = 0, \\ \bar{p}|_{t=T} = 0 \end{cases} \tag{5.1}$$

and

$$\operatorname{supp}\bar{\mu} \subset \bigl\{(x,t) \in Q \mid \bar{y}(x,t) = \varphi(x,t) \text{ or } \bar{y}(x,t) = \psi(x,t)\bigr\} \tag{5.2}$$

such that

$$H\bigl(x,t,\bar{y}(x,t),\bar{u}(x,t),\bar{p}(x,t)\bigr) = \min_{u \in U} H\bigl(x,t,\bar{y}(x,t),u,\bar{p}(x,t)\bigr)$$
$$\text{a.e. } (x,t) \in Q \tag{5.3}$$

where

$$H(x,t,y,u,p) = pf(x,t,y,u) + L(x,t,y,u)$$

for any $(x,t,y,u,p) \in \Omega \times [0,T] \times \mathbb{R} \times U \times \mathbb{R}$.

Remark. In the above, $\mathcal{M}_0(\overline{Q}) = C_0(\overline{Q})^*$ (where $C_0(\overline{Q}) = \{\eta \in C(\overline{Q}) | \eta|_\Sigma = 0\}$) is the set of all Radon measures on $\overline{Q}$ with the support contained in $Q \cup (\Omega \times \{0, T\})$. (5.1) and (5.3) are referred to as the adjoint equation (along the given optimal pair) and the Pontryagin's condition, respectively. The condition (5.2) is understood as the following: For any $\eta \in C_0(\overline{Q})$ with $\operatorname{supp}\eta \subset Q' = \{(x,t) \in Q \mid \varphi(x,t) < \bar{y}(x,t) < \psi(x,t)\}$,

$$\langle \bar{\mu}, \eta \rangle_{\mathcal{M}_0(\overline{Q}), C_0(\overline{Q})} = 0.$$

Proof. Let $(\bar{y}, \bar{u})$ be an optimal pair, given $\varepsilon > 0$ and

$$\alpha_\varepsilon = \bigl(J_\varepsilon(\bar{u}) - \bar{J}_\varepsilon + \varepsilon\bigr)^{1/2} > 0.$$

From Proposition 4.1(ii) and Theorem 4.6 we see that

$$J_\varepsilon(\bar{u}) - \bar{J}_\varepsilon \to J(\bar{u}) - \bar{J} = 0 \quad (\varepsilon \to 0)$$

and therefore

$$\alpha_\varepsilon \to 0 \quad (\varepsilon \to 0). \tag{5.4}$$

Since $J_\varepsilon(u)$ is continuous on $(\mathcal{U}, d)$ and

$$J_\varepsilon(\bar{u}) \leqslant \bar{J}_\varepsilon + \alpha_\varepsilon^2 = \inf_{u\in\mathcal{U}} J_\varepsilon(u) + \alpha_\varepsilon^2,$$

by Ekeland's variational principle there exists a $u_\varepsilon \in \mathcal{U}$, such that

$$d(u_\varepsilon, \bar{u}) \leqslant \alpha_\varepsilon, \tag{5.5}$$

$$-\alpha_\varepsilon d(u, u_\varepsilon) \leqslant J_\varepsilon(u) - J_\varepsilon(u_\varepsilon) \quad \forall u \in \mathcal{U}. \tag{5.6}$$

Let $(y_\varepsilon, u_\varepsilon) \in \mathcal{F}_\varepsilon$ and $v \in \mathcal{U}$ be fixed. By Proposition 4.2 we know that, for any $\rho \in (0,1)$, there exists a measurable set $E^\rho \subset Q$ with $m(E^\rho) = \rho m(Q)$, such that if we define

$$u_\varepsilon^\rho(x,t) = \begin{cases} u_\varepsilon(x,t) & \text{if } (x,t) \in Q \setminus E^\rho, \\ v(x,t) & \text{if } (x,t) \in E^\rho, \end{cases}$$

and let $(y_\varepsilon^\rho, u_\varepsilon^\rho) \in \mathcal{F}_\varepsilon$, then

$$\begin{cases} y_\varepsilon^\rho = y_\varepsilon + \rho z_\varepsilon + r_\varepsilon^\rho, \\ J_\varepsilon(u_\varepsilon^\rho) = J_\varepsilon(u_\varepsilon) + \rho j_\varepsilon + e_\varepsilon^\rho, \end{cases}$$

where z_ε and j_ε satisfy the following:

$$\begin{cases} z_{\varepsilon t} - \Delta z_\varepsilon + \left\{\frac{1}{\varepsilon}\left[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\right] - f_y(x,t,y_\varepsilon,u_\varepsilon)\right\} z_\varepsilon \\ \quad = f(x,t,y_\varepsilon,v) - f(x,t,y_\varepsilon,u_\varepsilon) \quad \text{in } Q, \\ z_\varepsilon|_{\partial_p Q} = 0 \end{cases} \tag{5.7}$$

and

$$j_\varepsilon = \int_Q \left[L_y(x,t,y_\varepsilon,u_\varepsilon) z_\varepsilon + L(x,t,y_\varepsilon,v) - L(x,t,y_\varepsilon,u_\varepsilon)\right] dx\,dt \tag{5.8}$$

with

$$\lim_{\rho\to 0} \frac{1}{\rho} \left\| r_\varepsilon^\rho \right\|_{C^{\theta,\theta/2}(\bar{Q})} = \lim_{\rho\to 0} \frac{1}{\rho} \left| e_\varepsilon^\rho \right| = 0$$

for some $\theta \in (0,1)$.

Now we take $u = u_\varepsilon^\rho$ in (5.6). It follows that

$$-\alpha_\varepsilon m(Q) \leqslant \frac{1}{\rho}\left[J_\varepsilon\left(u_\varepsilon^\rho\right) - J_\varepsilon(u_\varepsilon)\right] \to j_\varepsilon \quad (\rho \to 0). \tag{5.9}$$

Let $p_\varepsilon \in W$ be the unique solution of the following equation:

$$\begin{cases} -p_{\varepsilon t} - \Delta p_\varepsilon + \left\{\frac{1}{\varepsilon}\left[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\right] - f_y(x,t,y_\varepsilon,u_\varepsilon)\right\} p_\varepsilon \\ \quad = L_y(x,t,y_\varepsilon,u_\varepsilon) \quad \text{in } Q, \\ p_\varepsilon|_\Sigma = 0, \\ p_\varepsilon|_{t=T} = 0. \end{cases} \tag{5.10}$$

Then we may deduce from (5.7)–(5.10) that

$$\int_Q \big[H(x,t,y_\varepsilon,v,p_\varepsilon) - H(x,t,y_\varepsilon,u_\varepsilon,p_\varepsilon)\big]\,dx\,dt \geqslant -\alpha_\varepsilon m(Q). \tag{5.11}$$

In what follows, we are going to take the limits to get the final result.

First, by (5.4), (5.5), and Proposition 4.5 we have, for some $\theta \in (0,1)$,

$$\lim_{\varepsilon\to 0} \|y_\varepsilon - \bar{y}\|_{C^{\theta,\theta/2}(\overline{Q})\cap L^2(0,T;H_0^1(\Omega))} = 0. \tag{5.12}$$

Noting that $\beta' \geqslant 0$, $\gamma' \geqslant 0$, and $f_y \leqslant 0$ (cf. (2.14)), we can easily get the following estimate from Eq. (5.10):

$$\|p_\varepsilon\|_{L^2(0,T;H_0^1(\Omega))} \leqslant C. \tag{5.13}$$

Moreover, let $S_\delta(\cdot) \in C^1(\mathbb{R})$ be a family of smooth approximation to the sign function, satisfying the following:

$$S'_\delta(r) \geqslant 0, \quad \forall r \in \mathbb{R}$$

and

$$S_\delta(r) = \begin{cases} 1 & \text{if } r > \delta, \\ 0 & \text{if } r = 0, \\ -1 & \text{if } r < -\delta. \end{cases}$$

Multiplying Eq. (5.10) by $\varepsilon S_\delta(p_\varepsilon)$ and integrating it over Q we can get

$$\int_Q \big[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\big] p_\varepsilon S_\delta(p_\varepsilon)\,dx\,dt \leqslant C\varepsilon.$$

Letting $\delta \to 0$ we have

$$\big\|\big[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\big] p_\varepsilon\big\|_{L^1(Q)} \leqslant C\varepsilon. \tag{5.14}$$

In estimates (5.13) and (5.14), the constant C is independent of $\varepsilon > 0$. Hence we may let, extracting some subsequence if necessary,

$$\begin{cases} p_\varepsilon \to \bar{p} \quad \text{weakly in } L^2\big(0,T;H_0^1(\Omega)\big), \\ \frac{1}{\varepsilon}\big[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\big] p_\varepsilon \to \bar{\mu} \quad \text{weakly star in } \mathcal{M}_0(\overline{Q}). \end{cases}$$

Passing to the limit in (5.10) and (5.11), we obtain (5.1) and

$$\int_Q H\big(x,t,\bar{y}(x,t),v(x,t),\bar{p}(x,t)\big)\,dx\,dt \geqslant \int_Q H\big(x,t,\bar{y}(x,t),\bar{u}(x,t),\bar{p}(x,t)\big)\,dx\,dt \quad \forall v \in \mathcal{U}. \tag{5.15}$$

Consequently, Pontryagin's condition (5.3) follows from (5.15) in virtue of the separability of U and the continuity of the Hamiltonian H in the variable v.

For any $\eta \in C_0(\overline{Q})$ with $\operatorname{supp}\eta \subset Q' = \{(x,t) \in Q \mid \varphi(x,t) < \bar{y}(x,t) < \psi(x,t)\}$, the convergence (5.12) combined with the compactness of $\operatorname{supp}\eta$ ensures that, for some $\varepsilon_0 > 0$,

$$\varphi(x,t) < y_\varepsilon(x,t) < \psi(x,t) \quad \forall (x,t) \in \operatorname{supp}\eta,\ 0 < \varepsilon < \varepsilon_0,$$

which yields

$$\begin{aligned}\langle\bar{\mu}, \eta\rangle_{\mathcal{M}_0(\overline{Q}),C_0(\overline{Q})} &= \lim_{\varepsilon\to 0}\int\limits_Q \frac{1}{\varepsilon}\big[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\big]p_\varepsilon\eta\,dx\,dt \\ &= \lim_{\varepsilon\to 0}\int\limits_{\operatorname{supp}\eta} \frac{1}{\varepsilon}\big[\beta'(y_\varepsilon - \varphi) + \gamma'(y_\varepsilon - \psi)\big]p_\varepsilon\eta\,dx\,dt = 0.\end{aligned}$$

Thus (5.2) holds. The proof is complete. □

Acknowledgment

The author is grateful to Professor Xunjing Li and Professor Jiongmin Yong for their instructive suggestions and helpful advice.

References

[1] D.R. Adams, S.M. Lenhart, J. Yong, Optimal control of the obstacle for an elliptic variational inequality, Appl. Math. Optim. 38 (1998) 121–140.
[2] V. Barbu, Optimal Control of Variational Inequalities, Pitman, London, 1984.
[3] V. Barbu, Analysis and Control of Nonlinear Infinite Dimensional Systems, Academic Press, New York, 1993.
[4] J.F. Bonnans, D. Tiba, Pontryagin's principle in the control of semilinear elliptic variational inequalities, Appl. Math. Optim. 23 (1991) 299–312.
[5] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control Optim. 35 (1997) 1297–1327.
[6] L. Cesari, Optimization Theory and Applications, Problems with Ordinary Differential Equations, Springer, New York, 1983.
[7] Q. Chen, Indirect obstacle control problem for semilinear elliptic variational inequalities, SIAM J. Control Optim. 38 (1999) 138–158.
[8] Q. Chen, Indirect obstacle minimax control for elliptic variational inequalities, J. Optim. Theory Appl. 110 (2001) 337–359.
[9] I. Ekeland, Nonconvex minimization problems, Bull. Amer. Math. Soc. (N.S.) 1 (1979) 443–474.
[10] A. Friedman, Variational Principles and Free-Boundary Problems, Wiley, New York, 1982.
[11] A. Friedman, Optimal control for variational inequalities, SIAM J. Control Optim. 24 (1986) 439–451.
[12] A. Friedman, Optimal control for parabolic variational inequalities, SIAM J. Control Optim. 25 (1987) 482–497.
[13] B. Hu, J. Yong, Pontryagin maximum principle for semilinear and quasilinear parabolic equations with pointwise state constraints, SIAM J. Control Optim. 33 (1995) 1857–1880.
[14] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, Academic Press, New York, 1980.
[15] O.A. Ladyzhenskaya, V.A. Solonnikov, N.N. Ural'ceva, Linear and Quasi-Linear Equations of Parabolic Type, American Mathematical Society, Providence, RI, 1968.

[16] X. Li, J. Yong, Necessary conditions for optimal control of distributed parameter systems, SIAM J. Control Optim. 29 (1991) 895–908.
[17] X. Li, J. Yong, Optimal Control Theory for Infinite Dimensional Systems, Birkhäuser, Boston, 1995.
[18] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer, New York, 1971.
[19] F. Mignot, J.P. Puel, Optimal control in some variational inequalities, SIAM J. Control Optim. 22 (1984) 466–476.
[20] J.F. Rodrigues, Obstacle Problems in Mathematical Physics, in: North-Holland Mathematics Studies, Vol. 134, Elsevier, Amsterdam, 1987.
[21] J. Simon, Compact sets in the space $L^p(0,T;B)$, Ann. Mat. Pura Appl. 196 (1987) 65–96.
[22] R. Temam, Infinite-Dimensional Dynamical Systems in Mechanics and Physics, Springer, New York, 1988.
[23] G.M. Troianiello, Elliptic Differential Equations and Obstacle Problems, Plenum, New York, 1987.
[24] G. Wang, Optimal control of parabolic differential equations with two point boundary state constraints, SIAM J. Control Optim. 38 (2000) 466–476.
[25] J. Yong, Pontryagin maximum principle for semilinear second order elliptic partial differential equations and variational inequalities with state constraints, Differential Integral Equations 5 (1992) 1307–1334.
[26] J. Yong, Optimal control for nonlinear abstract evolution systems, Differential Integral Equations 6 (1993) 1145–1159.